ARMY RESEARCH LABORATORY

X-Wray Stats and Performance EXplorer

by Justin Wray, Carlos Mateo, Travis Parker, Ralph Ritchey, and Sidney Smith

ARL-TR-6895 April 2014

Approved for public release; distribution unlimited.

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory
Aberdeen Proving Ground, MD 21005

ARL-TR-6895 **April 2014**

X-Wray Stats and Performance EXplorer

Justin Wray, Carlos Mateo, Travis Parker, and Ralph Ritchey
ICF International

Sidney Smith
Computational and Informational Sciences Directorate, ARL

Approved for public release; distribution unlimited.

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)* April 2014	**2. REPORT TYPE** Final	**3. DATES COVERED (From - To)** October 2009–September 2010
4. TITLE AND SUBTITLE X-Wray Stats and Performance EXplorer		**5a. CONTRACT NUMBER**
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Justin Wray, Carlos Mateo, Travis Parker, Ralph Ritchey, and Sidney Smith		**5d. PROJECT NUMBER**
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory ATTN: RDRL-CIN-S Aberdeen Proving Ground, MD 21005		**8. PERFORMING ORGANIZATION REPORT NUMBER** ARL-TR-6895
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		**10. SPONSOR/MONITOR'S ACRONYM(S)**
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The X-Wray Stats and Performance EXplorer (X-Wray SPEX) is a test bed that allows intrusion detection tools to be rapidly tested, analyzed, and compared. The test bed itself consists of a small cluster employing the open grid engine to allow the automation of the testing process. Performance information (i.e., elapsed time, central processing unit load, memory usage, input/output [I/O] used, I/O wait, maximum virtual memory) is collected. In addition, X-Wray SPEX allows for large datasets to be thoroughly and quickly tagged by human analysts. The tool's output is compared to the human-based tagging to provide information about false positive and false negative detection rates. This report describes the test bed architecture and the tagging process and format, gives examples of the output format, and provides the results of the gage repeatability and reproducibility study conducted to validate the measurement capability of the test bed.

15. SUBJECT TERMS
measurement, performance, experimentation, security, verification

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Sidney Smith
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	UU	22	**19b. TELEPHONE NUMBER** (*Include area code*) 410-278-6235

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

1. Introduction

It is vitally important that intrusion detection tools be not only effective, but also efficient. The X-Wray Stats and Performance Explorer (X-Wray SPEX) test bed was developed to compare both the effectiveness and the efficiency of the locally developed X-Wray intrusion detection tool against the de facto standard intrusion detection tool Snort. X-Wray executes the tools to be evaluated against several datasets, including publicly available datasets like the Defense Advanced Research Projects Agency (DARPA) 1999,[1] Cyber Defense Exercise (CDX) 2009, Collegiate Cyber Defense Competition along with data captured in support of the U.S. Army Research Laboratory Computer Network Defense Service Provider that have been exhaustively tagged by network security analysts. We are able to compare not only the results for the tool against the tagging to calculate false positive and false negative rates (i.e., effectiveness), but also performance measures like CPU time and memory usage (i.e., efficiency).

1.1 Background

The effectiveness of intrusion detection systems (IDSs) is critically important. There have been several studies to compare the effectiveness of different IDS tools and techniques; however, there appears to have been very little work done on evaluating the efficiencies of IDS. The efficiency of IDS is critical because even a very effective IDS that is resource hungry may cause packet loss, preventing the engine from seeing malicious traffic. Schaelicke and Freeland observed packet loss rates of about 30% using Snort on commodity hardware.[2]

The analysis of the effectiveness of signature-based IDS turns out to be an evaluation of the rule set more than an evaluation of the engine itself. The very same engine with different rule sets will provide vastly different false positive and negative results. When comparing the efficiency of an IDS tool, one must tune the rule sets to provide the same detection results.

The X-Wray intrusion detection tool exploits the powerful regular expression-matching capabilities of the PERL language. We expected that regular expression-based rules would be more flexible and easier to create than the default rule format for Snort. Since Snort has grown quite large with many features added over the years, we also expected that a simpler, smaller engine would be more efficient. We have had experience with a simpler, more streamlined traffic capture program over the Tcpdump program. In order to compare the efficiencies of these tools, we tuned the rule sets until they provided almost identical results. X-Wray ran about 10 times

[1]Lippmann, R.; Fried, D.; Graf, I.; Haines, J.; Kendall, K.; McClung, D.; Weber, D.; Webster, S. E.; Wyschogrod, D.; Cunningham, R. K.; Zissman, M. A. Evaluating Intrusion Detection Systems: The 1998 DARPA Off-Line Intrusion Detection Evaluation. *DARPA Information Survivability Conference and Exposition* (DISCEX); Hilton Head, SC 2000.

[2]Schaelicke, L.; Freeland, J. C. Characterizing Sources and Remedies for Packet Loss in Network Intrusion Detection Systems. *Proceedings of the IEEE International Workload Characterization Symposium*; IEEE Conference Publications: Austin, TX, 2005.

longer than Snort primarily because of overhead associated with PERL and PERL's implementation of regular expressions.

1.2 Literature Review

Lincoln Laboratory, DARPA, and the U.S. Air Force worked together to evaluate IDS performance.[1] The focus of their work was the evaluation of IDS effectiveness. Perhaps the greatest contribution of that work was their off-line dataset. They created the test bed for the purpose of creating the off-line dataset. Instead of constructing a test bed to evaluate the systems, they released the dataset to the participants and collected their results for analysis.[1] This approach solved many of the problems inherent in the evaluation; however, it made evaluation of efficiency impossible.

1.3 Dataset Tagging

The Lincoln Laboratory project tagged their data with list files. The list files tagged each session recording a unique session ID, start date, start time, duration, service name, source port, destination port, source internet protocol (IP) address, destination IP address, score, and attack name.[1]

Sperotto et al.[3] created a labeled dataset for flow-based intrusion detection by creating a honey pot, collecting flow data, correlating that log data, and creating alerts. The dataset consisted of this flow data, log data, and alert data as related tables in a structure query language database—specifically, MySQL—with the following:

F = (Isrc, Idst, Psrc, Pdst, Pckts, Octs, Tstart, Tend, Flags, Prot)

L = (T, Isrc, Psrc, Idst, Pdst, Descr, Auto, Succ, Corr)

A = (T, Desr, Auto, Succ, Serv, Type)

2. Test Bed Architecture

The test bed architecture consists of a mini-cluster with one master node, three compute nodes, one shared data server, and a shared database server, as seen in figure 1. All compute nodes have access to the shared data, which are copied locally for processing to eliminate the artificial introduction of network latency and contention into the performance statistics. The Master node schedules and monitors jobs. Each job is spawned simultaneously on all three computation nodes measuring variance. The resulting efficiencies and effectiveness are fed to the shared database for later analysis.

[3]Sperotto, A.; Sadre, R.; Vliet, F.; Pras, A. A Labeled Data Set For Flow-based Intrusion Detection. *Proceedings of the 9th IEEE International Workshop on IP Operations and Management*; Springer-Verlag: Berlin, Germany, 2009.

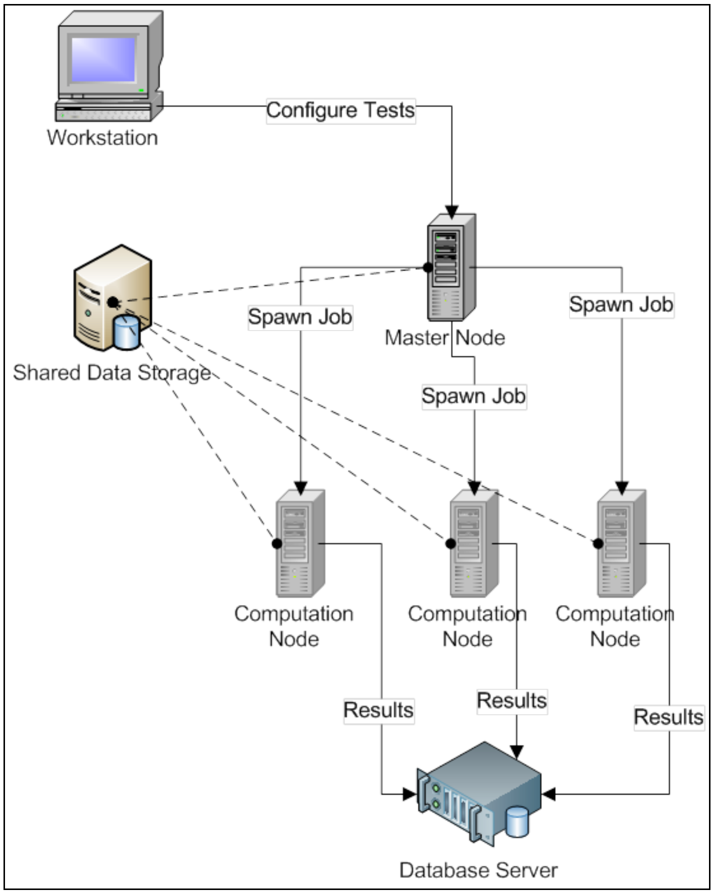

Figure 1. X-Wing logical architecture.

The information in table 1 is collected for each test conducted in the X-Wray SPEX to facilitate the analysis of the efficiency of the tool.

Table 1. X-Wray statistics efficiency data collected.

Name	**Description**
Job ID	A unique identifier for the individual testing instance
Job Type	An identifier for the tool user for the test instance
Data Set	An identifier for the dataset that was used for this test instance
Data File	Datasets are split into numerous smaller files (typically by hour)
Assigned Node	The individual computation node responsible for this test instance
Time Start	Time the test began
Time Finished	Time the test completed
Elapsed Time	Total time for the tool to process the given test (clock time)
CPU	CPU load information (amount of CPU time utilized)
Memory	Amount of memory untilized for the test
I/O	Input/Output used for the test
I/O Wait	Delay caused by reading/writing to/from media
Max VMem	Max virtual memory (page/swap) untilized for the test
Detects	Number of detects/alerts generated by the given tool for the given test

The information in table 2 is collected for each alert or detect to facilitate the analysis of the effectiveness of the tool.

Table 2. X-Wray statistics effectiveness data collected.

Name	Description
Tagging	For tagging categories indicated by smaller colorized buttons
Playback	ASCII and Hex playback links
Job ID	Job from which this detect was derived
Tool	Tool from which this detect was derived
Date	Date of the alerted traffic
Hour	Hour of the alerted traffic
Min	Minute of the alerted traffic
Sec	Second of the alerted traffic
Usec	Unix seconds of the alerted traffic
Source IP	Source Address of the alerted traffic
Source Port	Source application address of the alerted traffic
Dest IP	Destination Address of the alerted traffic
Dest Port	Destination application address of the alerted traffic
Protocol	Protocol of the alerted traffic
Flags	Flags of the alerted traffic
Type	Type code for this alerted traffic
Code	IP code for this alerted traffic
Alert	Textual message from the tool
Data	Relevant data delivered by the tool

3. Dataset Tagging

In order to compare the effectiveness of tools, it is necessary to compare the tools results to the "truth." Finding the truth turns out to be an interesting problem. Of the datasets that we used, only the DARPA dataset came with the flows tagged so that we would know the truth. However, the age of the dataset presents a problem, as standards have changed significantly. For example, back in 1998 it was common for users to log onto UNIX systems with telnet, exposing their passwords in the clear. Today such practices are considered poor security and prohibited. This is only one example of why traffic that would have been considered benign in 1998 would be considered a problem in 2012. To tag these datasets, our expert security analysts reviewed each flow and each alert using various tools to determine the level of concern for each flow. The level of concern was divided into the four categories enumerated in table 3.

Table 3. Tagging levels of concern.

Category	Description
Red	Root and User compromise. This category includes any instance where the adversary gained unauthorized access to the computer system either at the administration or user level.
Yellow	This category includes: • Attempted intrusions • Denial of service attacks • Poor security practices • Unauthorized scans • Malicious code
Green	Benign traffic
Blue	False positive

During this exercise, we found the definition of a false positive to be more interesting than originally anticipated. A false positive could be any of the following:

- When a rule fires, but the attack failed.
- When a rule fires, but the activity was not malicious.
- When a rule fires, but the activity doesn't really match the rule.

Each of these definitions is useful in a certain context; however, we needed to focus on what we were actually testing. The first two definitions would be excellent if we wanted to evaluate the rule set; however, we were evaluating the engine and therefore chose to use the third definition. We were surprised to discover that this was the case about 0.1% of the time.

4. Measurements

In order to ensure the validity of our measurement system, we conducted a Gage Reproducibility and Repeatability (Gage R&R) study. For the purposes of this study we consider each compute node to be an operator. We conducted the Long Form study with 10 samples measured three times each by three different operators for six different variables: Max VMem, CPU, detects, elapsed time, I/O, and memory. For this test, we took a publicly available rule set for Snort. X-Wray has a feature that allows it to import Snort rules. Since both engines used the same rule set, we were able to compare only the engines themselves. We took the data from table 4 and fed into it MiniTab for the statistical analysis. Reviewing the results, we found that the variation between parts was significantly larger than the variation between nodes or trials; therefore, we concluded that the measurement system is accurate enough for our purposes. Table 4 provides the Gage R&R data, and figures 2–7 show the Gage R&R results.

Table 4. Data for Gage R&R study.

Job ID	Job Type	Data File	Node	Elapsed Time	CPU	Memory	I/O	Max VMem	Detects
69815	snort	cdx010_20090421.09	xwray2	0.00	2.79	0.085	0.01	67.961	0
69816	snort	cdx010_20090421.09	xwray3	0.00	2.794	0.085	0.01	68.09	0
69817	snort	cdx010_20090421.09	xwray4	0.00	2.766	0.087	0.01	68.09	0
69805	xwray	cdx010_20090421.09	xwray2	0.00	161.646	5.744	0.01	44.973	5
69806	xwray	cdx010_20090421.09	xwray3	0.00	161.992	5.752	0.01	36.473	5
69807	xwray	cdx010_20090421.09	xwray4	0.00	163.173	5.793	0.01	44.973	5
70839	snort	cdx010_20090421.09	xwray2	0.00	2.803	0.14	0.01	68.086	0
70840	snort	cdx010_20090421.09	xwray3	0.00	2.786	0.155	0.01	68.09	0
70841	snort	cdx010_20090421.09	xwray4	0.00	2.775	0.144	0.01	68.09	0
69808	xwray	cdx010_20090421.09	xwray2	0.00	161.923	5.739	0.01	44.973	5
69809	xwray	cdx010_20090421.09	xwray3	0.00	161.934	5.754	0.01	36.473	5
69810	xwray	cdx010_20090421.09	xwray4	0.00	162.078	5.752	0.01	44.973	5
70842	snort	cdx010_20090421.09	xwray2	0.00	2.813	0.14	0.01	68.086	0
70843	snort	cdx010_20090421.09	xwray3	0.00	2.782	0.155	0.01	68.09	0
70844	snort	cdx010_20090421.09	xwray4	0.00	2.775	0.143	0.01	68.09	0
69811	xwray	cdx010_20090421.09	xwray2	0.00	162.554	5.745	0.01	44.973	5
69812	xwray	cdx010_20090421.09	xwray3	0.00	161.403	5.728	0.01	44.973	5
69813	xwray	cdx010_20090421.09	xwray4	0.00	161.933	5.747	0.01	44.973	5
69821	snort	cdx020_20090424.10	xwray2	0.00	49.773	2.825	0.33	69.25	6
69822	snort	cdx020_20090424.10	xwray3	0.00	49.87	2.878	0.33	69.258	6
69823	snort	cdx020_20090424.10	xwray4	0.00	49.548	2.824	0.33	69.254	6
69824	xwray	cdx020_20090424.10	xwray2	0.05	4282.35	152.484	0.33	44.973	9
69825	xwray	cdx020_20090424.10	xwray3	0.05	4284.14	152.564	0.33	36.473	9
69826	xwray	cdx020_20090424.10	xwray4	0.05	4279.95	152.422	0.33	44.973	9
70845	snort	cdx020_20090424.10	xwray2	0.00	49.464	2.821	0.33	60.75	6
70846	snort	cdx020_20090424.10	xwray3	0.00	49.92	2.88	0.33	69.258	6
70847	snort	cdx020_20090424.10	xwray4	0.00	49.383	2.816	0.33	69.258	6
69827	xwray	cdx020_20090424.10	xwray2	0.05	4300.51	153.128	0.33	44.973	9
69828	xwray	cdx020_20090424.10	xwray3	0.05	4285.52	152.615	0.33	44.973	9
69829	xwray	cdx020_20090424.10	xwray4	0.05	4327.57	154.119	0.33	44.973	9
70848	snort	cdx020_20090424.10	xwray2	0.00	49.76	2.877	0.33	69.25	6
70849	snort	cdx020_20090424.10	xwray3	0.00	49.758	2.878	0.33	69.262	6
70850	snort	cdx020_20090424.10	xwray4	0.00	49.366	2.815	0.33	69.254	6
69830	xwray	cdx020_20090424.10	xwray2	0.05	4285.23	152.583	0.33	44.973	9
69831	xwray	cdx020_20090424.10	xwray3	0.05	4286.89	152.662	0.33	44.973	9
69832	xwray	cdx020_20090424.10	xwray4	0.05	4294.83	152.937	0.33	36.473	9
69833	snort	darpa98_test_wk02_tue	xwray2	0.00	229.001	14.049	0.42	72.051	661
69834	snort	darpa98_test_wk02_tue	xwray3	0.00	228.433	14.026	0.42	72.059	661
69835	snort	darpa98_test_wk02_tue	xwray4	0.00	227.287	13.97	0.42	63.562	661
69865	xwray	darpa98_test_wk02_tue	xwray2	0.36	11291.1	402.142	0.44	44.973	524
69866	xwray	darpa98_test_wk02_tue	xwray3	0.35	11296.8	402.346	0.44	36.473	525
69867	xwray	darpa98_test_wk02_tue	xwray4	0.37	11352.4	404.293	0.44	44.973	523
70859	snort	darpa98_test_wk02_tue	xwray2	0.00	229.009	14.042	0.42	63.559	661
70860	snort	darpa98_test_wk02_tue	xwray3	0.00	228.742	14.042	0.42	72.074	661
70861	snort	darpa98_test_wk02_tue	xwray4	0.00	227.621	13.978	0.42	72.062	661

Table 4. Data for Gage R&R study (continued).

Job ID	Job Type	Data File	Node	Elapsed Time	CPU	Memory	I/O	Max VMem	Detects
69868	xwray	darpa98_test_wk02_tue	xwray2	0.36	11401.7	406.074	0.44	44.973	524
69869	xwray	darpa98_test_wk02_tue	xwray3	0.36	11350.8	404.253	0.44	44.973	527
69870	xwray	darpa98_test_wk02_tue	xwray4	0.37	11324.6	403.314	0.44	44.973	526
70862	snort	darpa98_test_wk02_tue	xwray2	0.00	230.325	14.162	0.42	72.062	661
70863	snort	darpa98_test_wk02_tue	xwray3	0.00	229.007	14.039	0.42	72.074	661
70864	snort	darpa98_test_wk02_tue	xwray4	0.00	227.525	13.972	0.42	72.062	661
69871	xwray	darpa98_test_wk02_tue	xwray2	0.36	11313.7	402.944	0.44	44.973	526
69872	xwray	darpa98_test_wk02_tue	xwray3	0.37	12212.4	434.932	0.44	44.973	523
69873	xwray	darpa98_test_wk02_tue	xwray4	0.37	11357.1	404.474	0.44	44.973	524
69836	snort	darpa98_training_wk04_wen	xwray2	0.00	33.898	1.938	0.08	68.832	179
69837	snort	darpa98_training_wk04_wen	xwray3	0.00	33.784	1.881	0.07	68.82	179
69838	snort	darpa98_training_wk04_wen	xwray4	0.00	33.325	1.879	0.07	68.816	179
70065	xwray	darpa98_training_wk04_wen	xwray3	0.01	1094.14	38.931	0.08	44.973	177
70068	xwray	darpa98_training_wk04_wen	xwray2	0.01	1099.35	39.125	0.08	44.973	177
70081	xwray	darpa98_training_wk04_wen	xwray4	0.01	1092.37	38.883	0.08	44.973	178
70865	snort	darpa98_training_wk04_wen	xwray2	0.00	33.796	1.897	0.07	60.223	179
70866	snort	darpa98_training_wk04_wen	xwray3	0.00	33.964	1.923	0.07	68.82	179
70867	snort	darpa98_training_wk04_wen	xwray4	0.00	33.372	1.898	0.07	68.828	179
70066	xwray	darpa98_training_wk04_wen	xwray3	0.01	1089.14	38.756	0.08	44.973	178
70076	xwray	darpa98_training_wk04_wen	xwray2	0.01	1094.97	38.987	0.08	44.973	178
70079	xwray	darpa98_training_wk04_wen	xwray4	0.01	1089.07	38.76	0.08	44.973	177
70868	snort	darpa98_training_wk04_wen	xwray2	0.00	33.807	1.945	0.08	68.723	179
70869	snort	darpa98_training_wk04_wen	xwray3	0.00	33.821	1.926	0.08	68.82	179
70870	snort	darpa98_training_wk04_wen	xwray4	0.00	33.31	1.897	0.08	68.844	179
70077	xwray	darpa98_training_wk04_wen	xwray2	0.01	1092.48	38.891	0.08	44.973	178
70080	xwray	darpa98_training_wk04_wen	xwray3	0.01	1093.15	38.916	0.08	44.973	177
70083	xwray	darpa98_training_wk04_wen	xwray4	0.01	1091.72	38.856	0.15	44.973	178
69839	snort	darpa99_wk05_inside_fri	xwray2	0.01	746.826	46.597	2.04	72.723	420
69840	snort	darpa99_wk05_inside_fri	xwray3	0.01	745.236	46.603	1.02	72.812	420
69841	snort	darpa99_wk05_inside_fri	xwray4	0.01	745.551	46.534	1.02	72.723	420
70078	xwray	darpa99_wk05_inside_fri	xwray3	0.18	15788.6	562.328	1.02	44.973	459
70084	xwray	darpa99_wk05_inside_fri	xwray2	0.18	15817.9	563.361	1.02	44.973	456
70094	xwray	darpa99_wk05_inside_fri	xwray4	0.18	15900.3	566.31	1.02	44.973	458
70871	snort	darpa99_wk05_inside_fri	xwray2	0.01	746.515	46.604	1.02	64.211	420
70872	snort	darpa99_wk05_inside_fri	xwray3	0.01	752.159	46.988	1.02	72.73	420
70873	snort	darpa99_wk05_inside_fri	xwray4	0.01	744.071	46.455	1.02	72.723	420
70082	xwray	darpa99_wk05_inside_fri	xwray3	0.18	15794.7	562.549	1.02	44.973	458
70090	xwray	darpa99_wk05_inside_fri	xwray2	0.18	15812.6	563.179	1.02	44.973	456
70092	xwray	darpa99_wk05_inside_fri	xwray4	0.18	15864.5	565.031	1.02	44.973	455
70874	snort	darpa99_wk05_inside_fri	xwray2	0.01	751.428	46.913	1.02	72.711	420
70875	snort	darpa99_wk05_inside_fri	xwray3	0.01	746.252	46.607	1.02	72.723	420
70876	snort	darpa99_wk05_inside_fri	xwray4	0.01	745.478	46.572	1.02	72.719	420
70091	xwray	darpa99_wk05_inside_fri	xwray2	0.18	15778.5	561.956	1.02	44.973	460
70093	xwray	darpa99_wk05_inside_fri	xwray3	0.18	15835.2	563.968	1.02	44.973	459
70095	xwray	darpa99_wk05_inside_fri	xwray4	0.18	15792.9	562.48	1.02	44.973	458

Figure 2 shows the results of the Gage R&R study for maximum virtual memory. As we look at the components of variation, we see that the part-to-part variation is significantly larger than any other source of variation. When we look at the R Chart and Xbar Charts by Operator, we find that they are both out of control. When we look at the Max VMem by Operator chart, we find that each operator performed almost identically. All of these indicate that the measurement system is performing well, and that we may trust the X-Wray SPEX to measure maximum virtual memory accurately.

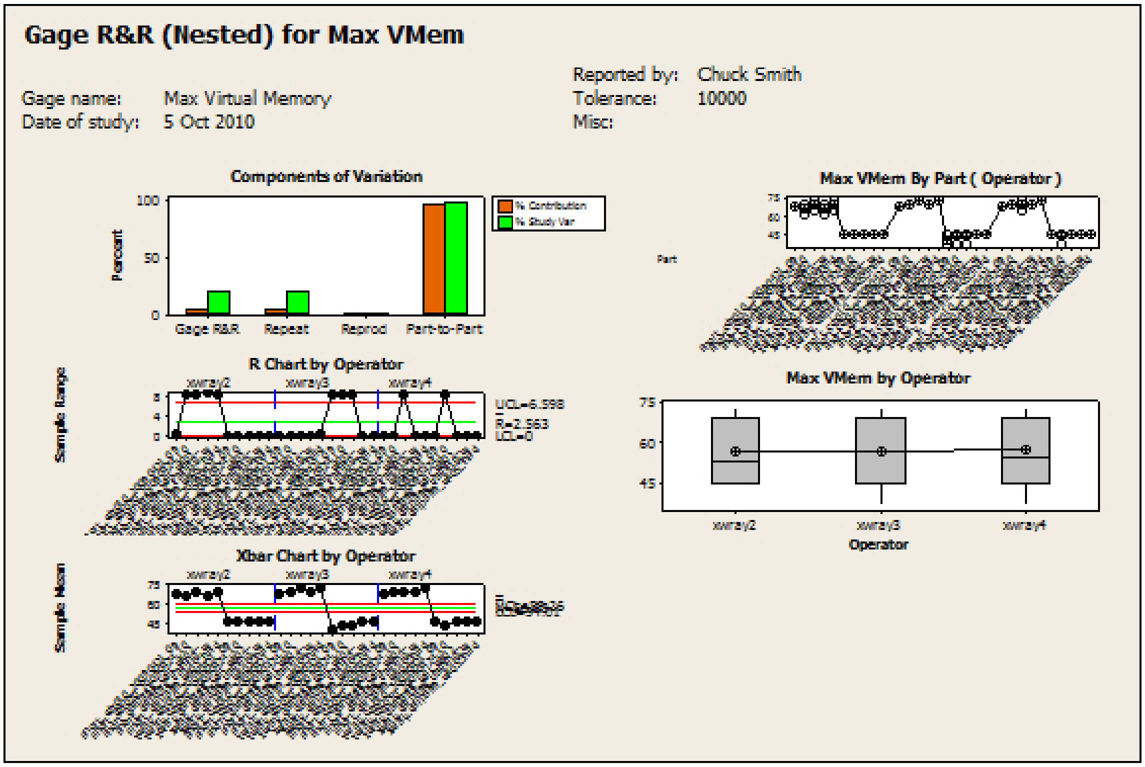

Figure 2. Gage R&R results for max VMem.

Figure 3 shows the results of the Gage R&R study for CPU cycles. As we look at the components of variation, we see that the part-to-part variation is significantly larger than any other source of variation. When we look at the R Chart and Xbar Charts by Operator, we find that they are both out of control. When we look at the CPU by Operator chart, we find that each operator performed almost identically. All of these indicate that the measurement system is performing well, and that we may trust the X-Wray SPEX to measure CPU cycles accurately.

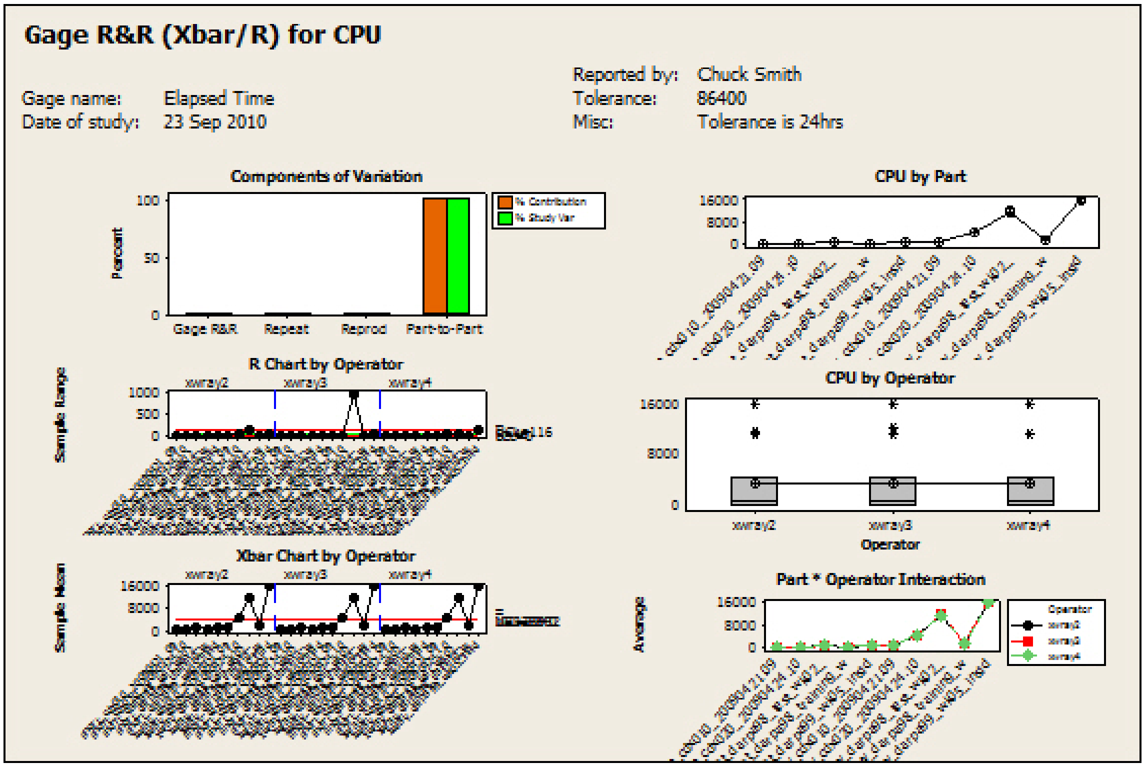

Figure 3. Gage R&R results for CPU.

Figure 4 shows the results of the Gage R&R study for detects. As we look at the components of variation, we see that the part-to-part variation is significantly larger than any other source of variation. When we look at the R Chart and Xbar Charts by Operator, we find that they are both out of control. When we look at the Detects by Operator chart, we find that each operator performed almost identically. All of these indicate that the measurement system is performing well, and that we may trust the X-Wray SPEX to measure detects accurately.

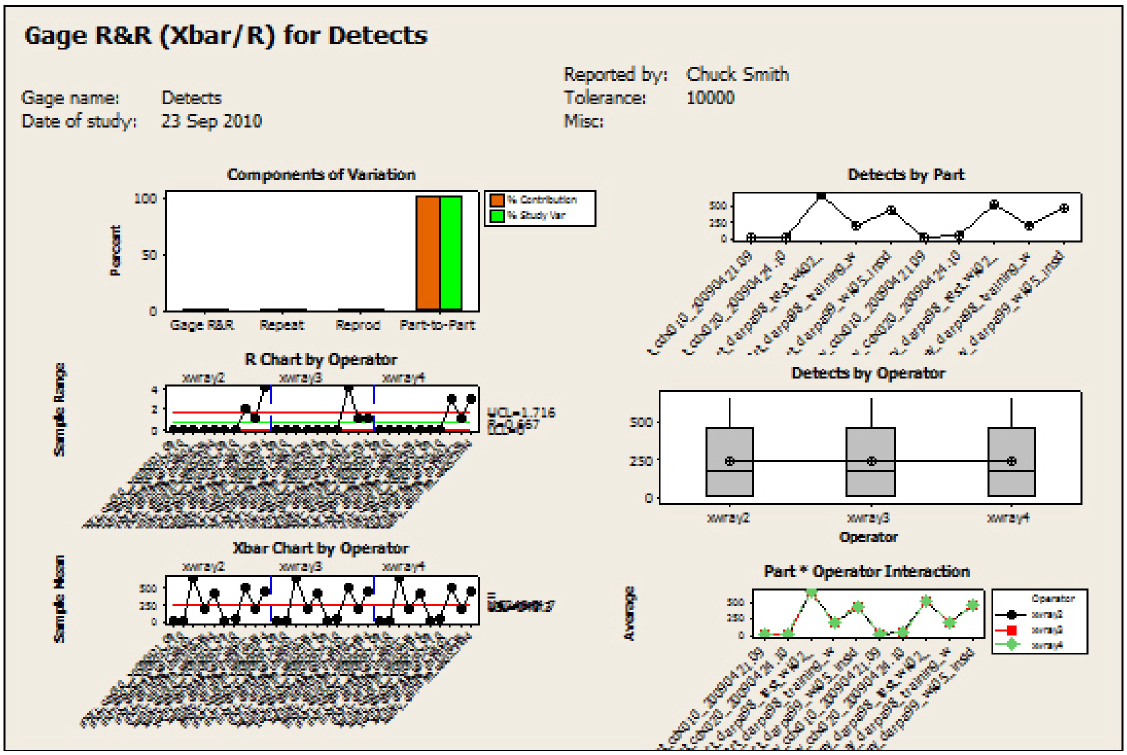

Figure 4. Gage R&R results for detects.

Figure 5 shows the results of the Gage R&R study for elapsed time. As we look at the components of variation, we see that the part-to-part variation is significantly larger than any other source of variation. When we look at the R Chart and Xbar Charts by Operator, we find that they are both out of control. When we look at the Elapsed Time by Operator chart, we find that each operator performed almost identically. All of these indicate that the measurement system is performing well, and that we may trust the X-Wray SPEX to measure elapsed time accurately.

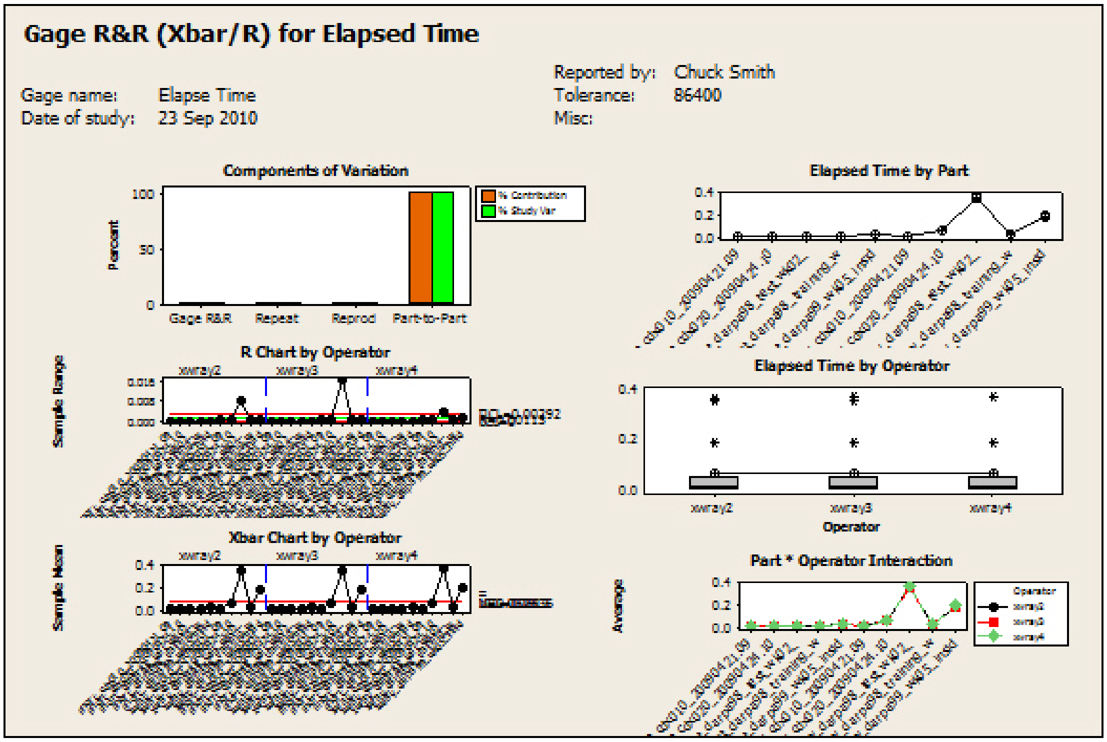

Figure 5. Gage R&R results for elapsed time.

Figure 6 shows the results of the Gage R&R study for I/O. As we look at the components of variation, we see that the part-to-part variation is significantly larger than any other source of variation. When we look at the R Chart and Xbar Charts by Operator, we find that they are both out of control. When we look at the I/O by Operator chart, we find that each operator performed almost identically. All of these indicate that the measurement system is performing well, and that we may trust the X-Wray SPEX to measure I/O accurately.

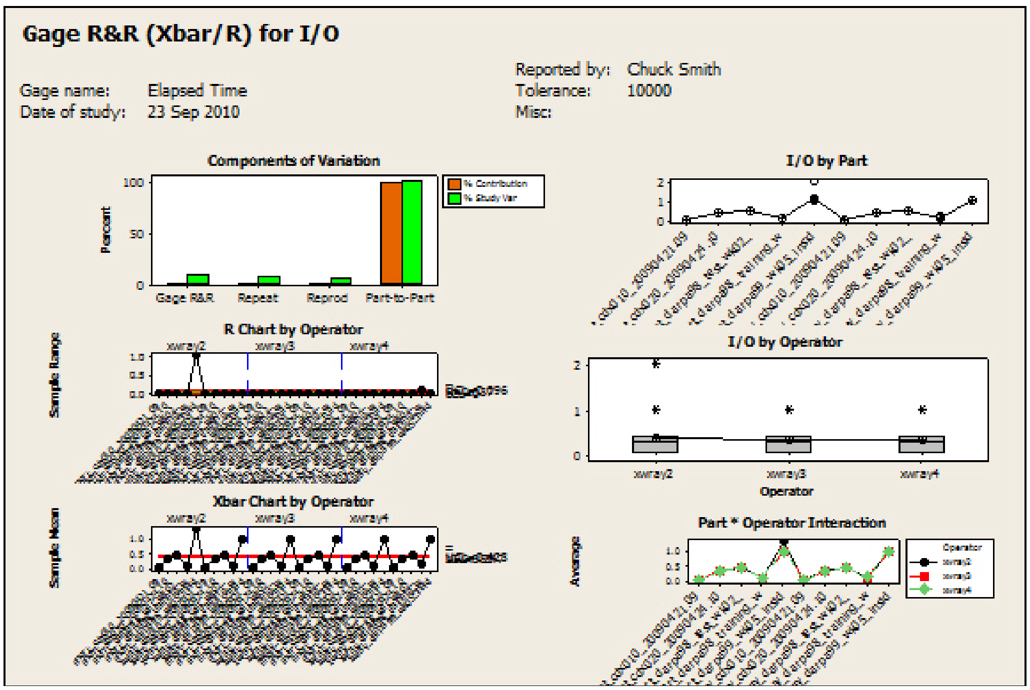

Figure 6. Gage R&R results for I/O.

Figure 7 shows the results of the Gage R&R study for memory. As we look at the components of variation, we see that the part-to-part variation is significantly larger than any other source of variation. When we look at the R Chart and Xbar Charts by Operator, we find that they are both out of control. When we look at the Memory by Operator chart, we find that each operator performed almost identically. All of these indicate that the measurement system is performing well, and that we may trust the X-Wray SPEX to measure memory accurately.

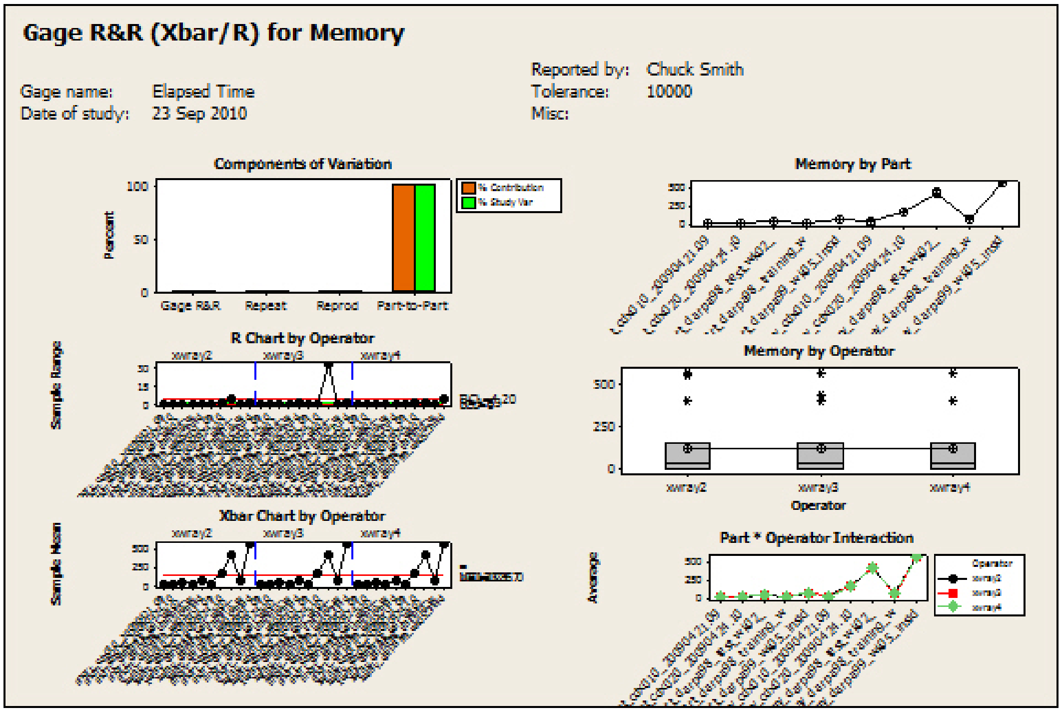

Figure 7. Gage R&R results for memory.

5. Conclusion

Although much work has been done to compare the effectiveness of network intrusion detection tools, little work has been done to compare the efficiency of these tools. X-Wray SPEX is a test bed designed to compare both the effectiveness and efficiency of these tools. We conducted a Gage R&R study to prove that the X-Wray SPEX test bed reliably measures key indicators of efficiency. We used this test bed to compare a locally developed tool to Snort, the de facto standard in network intrusion detection tools. The X-Wray SPEX was capable of clearly and accurately measuring the difference in the efficiency of these two tools.

NO. OF COPIES	ORGANIZATION
1 (PDF)	DEFENSE TECHNICAL INFORMATION CTR DTIC OCA
2 (PDF)	DIRECTOR US ARMY RESEARCH LAB RDRL CIO LL IMAL HRA MAIL & RECORDS MGMT
1 (PDF)	GOVT PRINTG OFC A MALHOTRA
1 (PDF)	DIR USARL RDRL CIN S S SMITH

INTENTIONALLY LEFT BLANK.